University of Bern
Institute of Mathematical Statistics and Actuarial Science
Technical Report 71

On an Auxiliary Function for Log-Density Estimation

Madeleine L. Cule and Lutz Dümbgen
(University of Cambridge and University of Bern)

July 2008

Abstract

In this note we provide explicit expressions and expansions for a special function J which appears in nonparametric estimation of log-densities. This function returns the integral of a log-linear function on a simplex of arbitrary dimension. In particular it is used in the R-package *LogCondDEAD* by Cule et al. (2007).

Contents

1 Introduction

Suppose one wants to estimate a probability density f on a certain compact region $C \subset \mathbb{R}^d$, based on an empirical distribution $\hat{P}$ of a sample from f. One possibility is to embed C into a union

$$S \;=\; \bigcup_{j=1}^{m} S_j$$

of simplices $S_j \subset \mathbb{R}^d$ with pairwise disjoint interior. By a simplex in $\mathbb{R}^d$ we mean the convex hull of $d+1$ points. Then we consider the family $\mathcal{G} = \mathcal{G}(S_1, \ldots, S_m)$ of all continuous functions $\psi : S \to \mathbb{R}$ which are linear on each simplex S_j. Now

$$\hat{\psi} \;:=\; \operatorname*{arg\,max}_{\psi \in \mathcal{G}} \left(\int_S \psi \, d\hat{P} - \int_S \exp(\psi(x))\, dx \right) \tag{1}$$

defines a maximum likelihood estimator $\hat{f} := \exp(\hat{\psi})$ of a probability density on S, based on $\hat{P}$. For existence and uniqueness of this estimator see, for instance, Cule et al. (2008).

To compute $\hat{\psi}$ explicitly, note that $\psi \in \mathcal{G}$ is uniquely determined by its values at the corners (extremal points) of all simplices S_j, and $\int \psi \, d\hat{P}$ is a linear function of these values. The second integral in (1) may be represented as follows: Let S_j be the convex hull of $\boldsymbol{x}_{0j}, \boldsymbol{x}_{1j}, \ldots, \boldsymbol{x}_{dj} \in \mathbb{R}^d$, and set $y_{ij} := \psi(\boldsymbol{x}_{ij})$. Then

$$\int_S \exp(\psi(x))\, dx \;=\; \sum_{i=1}^{m} \int_{S_i} \exp(\psi(x))\, dx \;=\; \sum_{i=1}^{m} D_j \cdot J(y_{0j}, y_{1j}, \ldots, y_{dj}),$$

where

$$D_j \ := \ \det\bigl[\boldsymbol{x}_{1j} - \boldsymbol{x}_{0j}, \boldsymbol{x}_{2j} - \boldsymbol{x}_{0j}, \ldots, \boldsymbol{x}_{dj} - \boldsymbol{x}_{0j}\bigr],$$

while $J(\cdot)$ is an auxiliary function defined and analyzed subsequently.

2 The special function $J(\cdot)$

2.1 Definition of $J(\cdot)$

For $d \in \mathbb{N}$ and $r \geq 0$ let

$$T_d(r) \ := \ \Bigl\{\boldsymbol{u} \in [0,\infty)^d : \sum_{i=1}^d u_i \leq r\Bigr\}.$$

Then for $y_0, y_1, \ldots, y_d \in \mathbb{R}$ we define

$$J(y_0, y_1, \ldots, y_d) \ := \ \int_{T_d(1)} \exp\Bigl((1-u_+)y_0 + \sum_{i=1}^d u_i y_i\Bigr)\, d\boldsymbol{u}$$

with $u_+ := \sum_{i=1}^d u_i$.

Standard considerations in connection with beta- and gamma-distributions as described in Section 6 reveal the following alternative representation:

$$J(y_0, y_1, \ldots, y_d) \ := \ \frac{1}{d!}\,\mathbb{E}\exp\Bigl(\sum_{i=0}^d B_i y_i\Bigr)$$

with $B_i = B_{d,i} := E_i \big/ \sum_{s=0}^d E_s$ and stochastically independent, standard exponential random variables $E_0, E_1, \ldots, E_d$. This representation shows clearly that $J(\cdot)$ is symmetric in its arguments.

An often useful identity is

$$J(y_0, y_1, \ldots, y_d) \ = \ \exp(y_*) J(y_0 - y_*, y_1 - y_*, \ldots, y_d - y_*) \quad \text{for any } y_* \in \mathbb{R}. \tag{2}$$

2.2 Some useful recursions

It is well-known that for any integer $0 \leq j < d$,

$$\Bigl(\frac{E_i}{\sum_{s=0}^j E_s}\Bigr)_{i=0}^j, \quad B := \frac{\sum_{i=0}^j E_i}{\sum_{s=0}^d E_s}, \quad \Bigl(\frac{E_i}{\sum_{s=j+1}^d E_s}\Bigr)_{i=j+1}^d$$

are stochastically independent with $B \sim \mathrm{Beta}(j+1, d-j)$; see also Section 6. Hence we end up with the following recursive identity:

$$\begin{aligned} & J(y_0, y_1, \ldots, y_d) \\ & \quad = \ \frac{j!(d-j-1)!}{d!}\,\mathbb{E}\bigl(J(By_0, \ldots, By_j) J((1-B)y_{j+1}, \ldots, (1-B)y_d)\bigr) \\ & \quad = \ \int_0^1 u^j (1-u)^{d-j-1} J(uy_0, \ldots, uy_j) J((1-u)y_{j+1}, \ldots, (1-u)y_d)\, du \end{aligned}$$

with

$$J(r) \ := \ \exp(r).$$

Here we used the well-known identity

$$\int (1-u)^\ell u^m \, du \ = \ \frac{\ell! m!}{(\ell+m+1)!} \quad \text{for integers } \ell, m \geq 0. \tag{3}$$

Plugging in $j = d-1$ into the previous recursive equation leads to

$$J(y_0, y_1, \ldots, y_d) \ = \ \int_0^1 u^{d-1} J(uy_0, \ldots, uy_{d-1}) \exp((1-u)y_d) \, du. \tag{4}$$

For $d = 1$ these recursive formulae are useless, but one compute $J(y_0, y_1)$ explicitly:

$$J(y_0, y_1) \ = \ \int_0^1 \exp\bigl((1-u)y_0 + uy_1\bigr) \, du \ = \ \begin{cases} \dfrac{\exp(y_1) - \exp(y_0)}{y_1 - y_0} & \text{if } y_0 \neq y_1, \\ \exp(y_0) & \text{if } y_0 = y_1. \end{cases}$$

But now we can prove an alternative recursive formula: For arbitrary integers $d \geq 1$ and real numbers $y_0, y_1, \ldots, y_d$,

$$J(y_0, y_1, \ldots, y_d) \ = \ \begin{cases} \dfrac{J(y_1, y_2, \ldots, y_d) - J(y_0, y_2, \ldots, y_d)}{y_1 - y_0} & \text{if } y_0 \neq y_1, \\ \dfrac{\partial}{\partial y_1} J(y_1, y_2, \ldots, y_d) & \text{if } y_0 = y_1. \end{cases} \tag{5}$$

(For $d = 1$, just erase "$y_2, \ldots$" on the right hand side.) We prove this formula by induction on d and restrict our attention to the case $y_0 \neq y_1$. The case $y_0 = y_1$ follows by a suitable limit argument. Suppose that (5) is true for some $d \geq 1$. Then it follows from (4), the induction hypothesis and a second application of (4) that

$$\begin{aligned}
&J(y_0, y_1, \ldots, y_d, y_{d+1}) \\
&= \ \int_0^1 u^d J(uy_0, uy_1, \ldots, uy_d) \exp((1-u)y_{d+1}) \, du \\
&= \ \int_0^1 u^d \, \frac{J(uy_1, uy_2, \ldots, uy_d) - J(uy_0, uy_2, \ldots, uy_d)}{uy_1 - uy_0} \exp((1-u)y_{d+1}) \, du \\
&= \ \int_0^1 u^{d-1} \, \frac{J(uy_1, uy_2, \ldots, uy_d) - J(uy_0, uy_2, \ldots, uy_d)}{y_1 - y_0} \exp((1-u)y_{d+1}) \, du \\
&= \ \frac{1}{y_1 - y_0} \Bigl(\int_0^1 u^{d-1} J(uy_1, uy_2, \ldots, uy_d) \exp((1-u)y_{d+1}) \, du \\
&\qquad\qquad - \int_0^1 u^{d-1} J(uy_0, uy_2, \ldots, uy_d) \exp((1-u)y_{d+1}) \, du \Bigr) \\
&= \ \frac{J(y_1, y_2, \ldots, y_d, y_{d+1}) - J(y_0, y_2, \ldots, y_d, y_{d+1})}{y_1 - y_0}.
\end{aligned}$$

3 An expansion for $J(\cdot)$

With $\bar{y} := (d+1)^{-1}\sum_{i=0}^d y_i$ and $z_i := y_i - \bar{y}$ one may write

$$J(y_0, y_1, \ldots, y_d) \; = \; \exp(\bar{y}) J(z_0, z_1, \ldots, z_d)$$

by virtue of (2). Note that $z_+ := \sum_{i=0}^d z_i = 0$. Thus, as $\boldsymbol{z} := (z_i)_{i=0}^d \to \mathbf{0}$,

$$\begin{aligned} d!\, J(z_0, z_1, \ldots, z_d) \; &= \; 1 + \sum_{i=0}^d \mathbb{E}(B_i) z_i + \frac{1}{2} \sum_{i,j=0}^d \mathbb{E}(B_i B_j) z_i z_j \\ &\quad + \frac{1}{6} \sum_{i,j,k=0}^d \mathbb{E}(B_i B_j B_k) z_i z_j z_k + O(\|\boldsymbol{z}\|^4). \end{aligned}$$

It follows from Lemma 6.1 that

$$\mathbb{E}\Bigl(\prod_{i=0}^d B_i^{k_i}\Bigr) \; = \; \prod_{i=0}^d k_i! \Big/ [d + k_+]_{k_+} \quad \text{for integers } k_0, k_1, \ldots, k_d \geq 0.$$

In particular,

$$\begin{aligned} \mathbb{E}(B_0) \; &= \; \frac{1}{d+1}, \\ \mathbb{E}(B_0^2) \; &= \; \frac{2}{[d+2]_2}, \qquad \mathbb{E}(B_0 B_1) \; = \; \frac{1}{[d+2]_2}, \\ \mathbb{E}(B_0^3) \; &= \; \frac{6}{[d+3]_3}, \qquad \mathbb{E}(B_0^2 B_1) \; = \; \frac{2}{[d+3]_3}, \qquad \mathbb{E}(B_0 B_1 B_2) \; = \; \frac{1}{[d+3]_3}. \end{aligned}$$

Consequently, $\sum_{i=0}^d \mathbb{E}(B_i) z_i = \mathbb{E}(B_0) z_+ = 0$,

$$\begin{aligned} [d+2]_2 \sum_{i,j=0}^d \mathbb{E}(B_i B_j) z_i z_j \; &= \; \sum_{i,j=0}^d \bigl(1\{i=j\} \cdot 2 + 1\{i \neq j\}\bigr) z_i z_j \\ &= \; \sum_{i,j=0}^d \bigl(1\{i=j\} + 1\bigr) z_i z_j \\ &= \; \sum_{i=0}^d z_i^2 + z_+^2 \\ &= \; \sum_{i=0}^d z_i^2, \end{aligned}$$

and

$$
\begin{aligned}
&[d+3]_3 \sum_{i,j,k=0}^{d} \mathbb{E}(B_i B_j B_k) z_i z_j z_k \\
&= \sum_{i,j,k=0}^{d} \big(1\{i=j=k\}\cdot 6 + 1\{\#\{i,j,k\}=2\}\cdot 2 + 1\{\#\{i,j,k\}=3\}\big) z_i z_j z_k \\
&= \sum_{i,j,k=0}^{d} \big(1\{i=j=k\}\cdot 5 + 1\{\#\{i,j,k\}=2\} + 1\big) z_i z_j z_k \\
&= 5\sum_{i=0}^{d} z_i^3 + 3\sum_{s,t=0}^{d} 1\{s\neq t\} z_s^2 z_t + z_+^3 \\
&= 5\sum_{i=0}^{d} z_i^3 + 3\sum_{s=0}^{d} z_s^2 z_+ - 3\sum_{s=0}^{d} z_s^3 + z_+^3 \\
&= 2\sum_{i=0}^{d} z_i^3.
\end{aligned}
$$

Consequently,

$$
J(y_0, y_1, \ldots, y_d) \;=\; \exp(\bar{y})\Big(\frac{1}{d!} + \frac{1}{2(d+2)!}\sum_{i=0}^{d} z_i^2 + \frac{1}{3(d+3)!}\sum_{i=0}^{d} z_i^3 + O\big(\|\boldsymbol{z}\|^4\big)\Big). \tag{6}
$$

4 A recursive implementation of $J(\cdot)$ and its partial derivatives

By means of (5) and the Taylor expansion (6) one can implement the function $J(\cdot)$ in a recursive fashion. In what follows we use the abbreviation

$$
y_{a:b} \;=\; \begin{cases} (y_a, \ldots, y_b) & \text{if } a \leq b \\ () & \text{if } a > b \end{cases}
$$

To compute $J(y_{0:d})$ we assume without loss of generality that $y_0 \leq y_1 \leq \cdots \leq y_d$. It follows from (5) and symmetry of $J(\cdot)$ that

$$
J(y_{0:d}) \;=\; \frac{J(y_{1:d}) - J(y_{0:d-1})}{y_d - y_0}
$$

if $y_0 \neq y_d$. This formula is okay numerically if $y_d - y_0$ is not too small. Otherwise one should use (6). This leads to the the pseudo code in Table 1.

To avoid messy formulae, one can express partial derivatives of $J(\cdot)$ in terms of higher order versions of $J(\cdot)$ by means of the recursion (5). For instance,

$$
\begin{aligned}
\frac{\partial J(y_{0:d})}{\partial y_0} &= \lim_{\epsilon\to 0} \frac{J(y_0+\epsilon,\, y_{1:d}) - J(y_0,\, y_{1:d})}{\epsilon} \\
&= \lim_{\epsilon\to 0} J(y_0, y_0+\epsilon,\, y_{1:d}) \\
&= J(y_0, y_0, y_{1:d}).
\end{aligned}
$$

Algorithm $J \leftarrow \mathbf{J}(y, d, \epsilon)$
if $y_d - y_0 < \epsilon$ **then**
$\quad\bar{y} \leftarrow \sum_{i=0}^{d} y_i/(d+1)$
$\quad z_2 \leftarrow \sum_{i=0}^{d}(y_i - \bar{y})^2/2$
$\quad z_3 \leftarrow \sum_{i=0}^{d}(y_i - \bar{y})^3/3$
$\quad J \leftarrow \exp(\bar{y})\big(1/d! + z_2/(d+2)! + z_3/(d+3)!\big)$
else
$\quad J \leftarrow \big(\mathbf{J}(y_{1:d}, d-1, \epsilon) - \mathbf{J}(y_{0:d-1}, d-1, \epsilon)\big)/(y_d - y_0)$
end if.

Table 1: Pseudo-code for $J(y)$ with ordered input vector y.

Similarly,

$$\begin{aligned}
\frac{\partial^2 J(y_{0:d})}{\partial y_0^2} &= \lim_{\epsilon \to 0}\Big(\frac{J(y_0+\epsilon,\, y_{1:d}) - J(y_0,\, y_{1:d})}{\epsilon} - \frac{J(y_0,\, y_{1:d}) - J(y_0-\epsilon,\, y_{1:d})}{\epsilon}\Big)/\epsilon \\
&= 2 \lim_{\epsilon \to 0} \frac{J(y_0, y_0+\epsilon,\, y_{1:d}) - J(y_0, y_0-\epsilon,\, y_{1:d})}{2\epsilon} \\
&= 2 \lim_{\epsilon \to 0} J(y_0, y_0-\epsilon, y_0+\epsilon,\, y_{1:d}) \\
&= 2\, J(y_0, y_0, y_0,\, y_{1:d}),
\end{aligned}$$

while

$$\begin{aligned}
\frac{\partial^2 J(y_{0:d})}{\partial y_0 \partial y_1} &= \lim_{\epsilon \to 0}\Big(\frac{J(y_0+\epsilon, y_1+\epsilon,\, y_{2:d}) - J(y_0, y_1+\epsilon,\, y_{2:d})}{\epsilon} \\
&\qquad\qquad - \frac{J(y_0+\epsilon, y_1,\, y_{2:d}) - J(y_0, y_1,\, y_{2:d})}{\epsilon}\Big)/\epsilon \\
&= \lim_{\epsilon \to 0} \frac{J(y_0, y_0+\epsilon, y_1+\epsilon,\, y_{2:d}) - J(y_0, y_0+\epsilon, y_1,\, y_{2:d})}{\epsilon} \\
&= \lim_{\epsilon \to 0} J(y_0, y_0+\epsilon, y_1, y_1+\epsilon,\, y_{2:d}) \\
&= J(y_0, y_0, y_1, y_1,\, y_{2:d}).
\end{aligned}$$

5 The special cases $d = 1$ and $d = 2$

For small dimension d it may be worthwhile to work with non-recursive implementations of the function $J(\cdot)$. Here we collect and extend some results of Dümbgen et al. (2007).

5.1 General considerations about a bivariate function

In view of (5) we consider an arbitrary function $f : \mathbb{R} \to \mathbb{R}$ which is infinitley often differentiable. Then

$$h(r,s) \;:=\; \begin{cases} \dfrac{f(s) - f(r)}{s - r} & \text{if } s \neq r, \\[2ex] f'(r) + \dfrac{f''(r)}{2}(s-r) + O\big((s-r)^2\big) & \text{as } s \to r, \end{cases}$$

defines a smooth and symmetric function $h : \mathbb{R}^2 \to \mathbb{R}$. Its first partial derivatives of order one and two are given by

$$\frac{\partial h(r,s)}{\partial r} \;=\; \begin{cases} \dfrac{f(s) - f(r) - f'(r)(s-r)}{(s-r)^2} & \text{if } s \neq r, \\ \dfrac{f''(r)}{2} + \dfrac{f'''(r)}{6}(s-r) + O\big((s-r)^2\big) & \text{as } s \to r, \end{cases}$$

$$\frac{\partial^2 h(r,s)}{\partial r^2} \;=\; \begin{cases} \dfrac{2\big(f(s) - f(r) - f'(r)(s-r)\big) - (s-r)^2 f''(r)}{(s-r)^3} & \text{if } s \neq r, \\ \dfrac{f'''(r)}{3} + \dfrac{f''''(r)}{12}(s-r) + O\big((s-r)^2\big) & \text{as } s \to r, \end{cases}$$

$$\frac{\partial^2 h(r,s)}{\partial r \partial s} \;=\; \begin{cases} \dfrac{(s-r)\big(f'(r) + f'(s)\big) - 2\big(f(s) - f(r)\big)}{(s-r)^3} & \text{if } s \neq r, \\ \dfrac{f'''(r)}{6} + \dfrac{f''''(r)}{12}(s-r) + O\big((s-r)^2\big) & \text{as } s \to r. \end{cases}$$

The other partial derivatives of order one and two follow via symmetry considerations.

5.2 More details for the case $d = 1$

Here the general formula boils down to

$$J(r,s) \;=\; \int_0^1 \exp\big((1-u)r + us\big)\, du,$$

and elementary calculations show that

$$J(r,s) \;=\; \begin{cases} \dfrac{\exp(s) - \exp(r)}{s - r} & \text{if } r \neq s, \\ \exp(r) & \text{if } r = s. \end{cases}$$

This is just the function introduced by Dümbgen, Hüsler and Rufibach (2007). Let us recall some properties and formulae for the corresponding partial derivatives

$$J_{a,b}(r,s) \;:=\; \frac{\partial^{a+b}}{\partial r^a \partial s^b}\, J(r,s) \;=\; \int_0^1 (1-u)^a u^b \exp((1-u)r + us)\, du.$$

Note first that

$$J_{a,b}(r,s) \;=\; J_{b,a}(s,r) \;=\; \exp(r) J_{a,b}(0, s-r).$$

Thus it suffices to derive formulae for $(r,s) = (0,y)$ and $b \le a$. It follows from (3) that

$$\begin{aligned}
J_{a,0}(0,y) &= \int_0^1 (1-u)^a \sum_{k=0}^{\infty} \frac{u^k}{k!}\, y^k \, du \\
&= \sum_{k=0}^{\infty} \frac{1}{k!} \int_0^1 (1-u)^a u^k \, du \cdot y^k \\
&= \sum_{k=0}^{\infty} \frac{a!}{(k+a+1)!}\, y^k \\
&= \frac{a!}{y^{a+1}} \Big(\exp(y) - \sum_{\ell=0}^{a} \frac{y^\ell}{\ell!} \Big).
\end{aligned}$$

In particular,

$$\begin{aligned}
J_{1,0}(0,y) &= \frac{\exp(y) - 1 - y}{y^2} \\
&= \frac{1}{2} + \frac{y}{6} + \frac{y^2}{24} + \frac{y^3}{120} + O(y^4) \quad (y \to 0), \\
J_{2,0}(0,y) &= \frac{2(\exp(y) - 1 - y - y^2/2)}{y^3} \\
&= \frac{1}{3} + \frac{y}{12} + \frac{y^2}{60} + \frac{y^3}{360} + O(y^4) \quad (y \to 0), \\
J_{3,0}(0,y) &= \frac{6(\exp(y) - 1 - y - y^2/2 - y^3/6)}{y^4} \\
&= \frac{1}{4} + \frac{y}{20} + \frac{y^2}{120} + \frac{y^3}{840} + O(y^4) \quad (y \to 0), \\
J_{4,0}(0,y) &= \frac{24(\exp(y) - 1 - y - y^2/2 - y^3/6 - y^4/24)}{y^5} \\
&= \frac{1}{5} + \frac{y}{30} + \frac{y^2}{210} + \frac{y^3}{1680} + O(y^4) \quad (y \to 0).
\end{aligned}$$

Another general observation is that

$$\begin{aligned}
J_{a,b}(r,s) &= \int_0^1 (1-u)^a (1-(1-u))^b \exp((1-u)r + us)\, du \\
&= \sum_{i=0}^{b} \binom{b}{i} (-1)^i J_{a+i,0}(r,s).
\end{aligned}$$

In particular,

$$\begin{aligned}
J_{a,1}(r,s) &= J_{a,0}(r,s) - J_{a+1,0}(r,s), \\
J_{a,2}(r,s) &= J_{a,0}(r,s) - 2J_{a+1,0}(r,s) + J_{a+2,0}(r,s).
\end{aligned}$$

On the other hand,

$$\begin{aligned} J_{a,b}(0,y) &= \sum_{k=0}^{\infty} \frac{y^k}{k!} \int_0^1 (1-u)^a u^{k+b}\, du \\ &= \sum_{k=0}^{\infty} \frac{a![k+b]_b}{(k+a+b+1)!}\, y^k \end{aligned}$$

with $[r]_0 := 1$ and $[r]_m := \prod_{i=0}^{m-1}(r-i)$ for integers $m > 0$. In particular,

$$\begin{aligned} J_{1,1}(0,y) &= \frac{\exp(y)(y-2)+2+y}{y^3} \\ &= \frac{1}{6} + \frac{y}{12} + \frac{y^2}{40} + \frac{y^3}{180} + O(y^4) \quad (y \to 0). \end{aligned}$$

5.3 The case $d = 2$

Our recursion formula (5) yields

$$J(r,s,t) = \begin{cases} \dfrac{J(s,t) - J(r,t)}{s-r} & \text{if } r \neq s, \\ J_{10}(r,t) & \text{if } r = s. \end{cases}$$

Because of J's symmetry we may rewrite this in terms of the order statistics $y_{(0)} \leq y_{(1)} \leq y_{(2)}$ of $(y_i)_{i=0}^2$ as

$$J(r,s,t) = \begin{cases} \dfrac{J(y_{(1)}, y_{(2)}) - J(y_{(0)}, y_{(1)})}{y_{(2)} - y_{(0)}} & \text{if } y_{(0)} < y_{(2)}, \\ \dfrac{\exp(y_{(0)})}{2} & \text{if } y_{(0)} = y_{(2)}. \end{cases}$$

For fixed third argument t, this function $J(r,s,t)$ corresponds to $h(r,s)$ in Section 5.1 with $f(x) := J(x,t)$. Thus

$$\frac{\partial J(r,s,t)}{\partial r} = \begin{cases} \dfrac{J(s,t) - J(r,t) - J_{1,0}(r,t)(s-r)}{(s-r)^2} & \text{if } r \neq s, \\ \dfrac{J_{2,0}(r,t)}{2} + \dfrac{J_{3,0}(r,t)(s-r)}{6} + O\big((s-r)^2\big) & \text{as } s \to r. \end{cases}$$

Moreover,

$$\frac{\partial^2 J(r,s,t)}{\partial r^2} = \begin{cases} \dfrac{2\big(J(s,t) - J(r,t) - J_{1,0}(r,t)(s-r)\big) - (s-r)^2 J_{2,0}}{(s-r)^3} & \text{if } r \neq s, \\ \dfrac{J_{3,0}(r,t)}{3} + \dfrac{J_{4,0}(r,t)(s-r)}{12} + O\big((s-r)^2\big) & \text{as } s \to r, \end{cases}$$

$$\frac{\partial^2 J(r,s,t)}{\partial r \partial s} = \begin{cases} \dfrac{\big(J_{1,0}(r,t) + J_{1,0}(s,t)\big)(s-r) - 2\big(J(s,t) - J(r,t)\big)}{(s-r)^3} & \text{if } r \neq s, \\ \dfrac{J_{3,0}(r,t)}{6} + \dfrac{J_{4,0}(r,t)(s-r)}{12} + O\big((s-r)^2\big) & \text{as } s \to r. \end{cases}$$

6 Gamma and multivariate beta (Dirichlet) distributions

Let $G_0, G_1, \ldots, G_m$ be stochastically independent random variables with $G_i \sim \mathrm{Gamma}(a_i)$ for certain parameters $a_i > 0$. That means, for any Borel set $A \subset (0,\infty)$,

$$\mathbb{P}(G_i \in A) \;=\; \int_A \Gamma(a_i)^{-1} y^{a_i - 1} \exp(-y)\, dy.$$

Now we define $a_+ := \sum_{i=0}^m a_i$, $G_+ := \sum_{i=0}^m G_i$ and

$$\tilde{\boldsymbol{B}} \;:=\; (G_i/G_+)_{i=0}^m, \quad \boldsymbol{B} \;:=\; (G_i/G_+)_{i=1}^m.$$

Note that $\tilde{\boldsymbol{B}}$ is contained in the unit simplex in $\mathbb{R}^{m+1}$, while $\boldsymbol{B}$ is contained in the set

$$\mathcal{T}_m \;:=\; \big\{\boldsymbol{u} \in (0,1)^m : u_+ < 1\big\}$$

with $u_+ := \sum_{i=1}^m u_i$. We also define $u_0 := 1 - u_+$ for any $\boldsymbol{u} \in \mathcal{T}_m$.

Lemma 6.1 *The random vector $\boldsymbol{B}$ and the random variable G_+ are stochastically independent. Moreover,*

$$G_+ \;\sim\; \mathrm{Gamma}(a_+)$$

while $\boldsymbol{B}$ is distributed according to the Lebesgue density

$$f(\boldsymbol{u}) \;:=\; \frac{\Gamma(a_+)}{\prod_{i=0}^m \Gamma(a_i)} \prod_{i=0}^m u_i^{a_i - 1}$$

on $\mathcal{T}_m$. For arbitrary numbers $k_0, k_1, \ldots, k_m \ge 0$ and $k_+ := \sum_{i=0}^m k_i$,

$$\mathbb{E}\Big(\prod_{i=0}^m B_i^{k_i}\Big) \;=\; \frac{\Gamma(a_+)}{\Gamma(a_+ + k_+)} \prod_{i=0}^m \frac{\Gamma(a_i + k_i)}{\Gamma(a_i)}.$$

As a by-product of this lemma we obtain the following formula:

Corollary 6.2 *For arbitrary numbers $a_0, a_1, \ldots, a_m > 0$,*

$$\int_{\mathcal{T}_m} \prod_{i=0}^m u_i^{a_i - 1}\, d\boldsymbol{u} \;=\; \Gamma(a_+)^{-1} \prod_{i=0}^m \Gamma(a_i).$$

Proof of Lemma 6.1. Note that $\boldsymbol{G} = (G_i)_{i=0}^m$ my be written as $\Xi(G_+, \boldsymbol{B})$ with the bijective mapping $\Xi : (0,\infty) \times \mathcal{T}_m \to (0,\infty)^{m+1}$,

$$\Xi(s, \boldsymbol{u}) \;:=\; (s u_i)_{i=0}^m.$$

Note also that

$$\det D\Xi(s,\boldsymbol{u}) \;=\; \det \begin{pmatrix} u_0 & -s & -s & \cdots & -s \\ u_1 & s & 0 & \cdots & 0 \\ u_2 & 0 & s & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ u_m & 0 & \cdots & 0 & s \end{pmatrix} \;=\; \det \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ u_1 & s & 0 & \cdots & 0 \\ u_2 & 0 & s & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ u_m & 0 & \cdots & 0 & s \end{pmatrix} \;=\; s^m.$$

Thus the distribution of $(G_+, \boldsymbol{B})$ has a Lebesgue density h on $(0,\infty) \times \mathcal{T}_m$ which is given by

$$\begin{aligned} h(s,\boldsymbol{u}) &= \prod_{i=0}^m \bigl(\Gamma(a_i)^{-1}\Xi(s,\boldsymbol{u})_i^{a_i-1}\exp(-\Xi(s,\boldsymbol{u})_i)\bigr) \cdot \bigl|\det D\Xi(s,\boldsymbol{u})\bigr| \\ &= \prod_{i=0}^m \bigl(\Gamma(a_i)^{-1}(su_i)^{a_i-1}\exp(-su_i)\bigr) \cdot s^m \\ &= s^{a_+-1}\exp(-s)\prod_{i=0}^m \bigl(\Gamma(a_i)^{-1}u_i^{a_i-1}\bigr) \\ &= \Gamma(a_+)^{-1}s^{a_+-1}\exp(-s)\cdot f(\boldsymbol{u}). \end{aligned}$$

Since this is the density of $\mathrm{Gamma}(a_+)$ at s times $f(\boldsymbol{u})$, we see that G_+ and $\boldsymbol{B}$ are stochastically independent, where G_+ has distribution $\mathrm{Gamma}(a_+)$, and that f is indeed a probability density on $\mathcal{T}_m$ describing the distribution of $\boldsymbol{B}$.

The fact that f integrates to one over $\mathcal{T}_m$ entails Corollary 6.2. But then we can conclude that

$$\begin{aligned} \mathbb{E}\Bigl(\prod_{i=0}^m B_i^{k(i)}\Bigr) &= \int_{\mathcal{T}_m}\prod_{i=0}^m u_i^{a_i+k_i-1}\,d\boldsymbol{u}\Big/\int_{\mathcal{T}_m}\prod_{i=0}^m u_i^{a_i-1}\,d\boldsymbol{u} \\ &= \frac{\Gamma(a_+)}{\Gamma(a_++k_+)}\prod_{i=0}^m\frac{\Gamma(a_i+k_i)}{\Gamma(a_i)}. \qquad \Box \end{aligned}$$

References

[1] M.L. CULE, R.B. GRAMACY, and R.J. SAMWORTH (2007). LogConcDEAD, An R package for log-concave density estimation in arbitrary dimensions. Available from `http://cran.r-project.org/`.

[2] M.L. CULE, R.J. SAMWORTH and M.I. STEWART (2008). Maximum likelihood estimation of a multidimensional log-concave density. Preprint.

[3] L. DÜMBGEN, A. HÜSLER and K. RUFIBACH (2007). Active set and EM algorithms for log-concave densities based on complete and censored data. Technical report 61, IMSV, University of Bern (arXiv:0707.4643).